UNPUBLISHED PRELIMINARY DATA

[redacted]

GPO PRICE $ ________

OTS PRICE(S) $ ________

Hard copy (HC) $ 2.00

Microfiche (MF) .50

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER

COLLEGE PARK, MARYLAND

Technical Report TR-65-14
NsG-398 February 1965

ON A GENERAL ESTIMATION PRINCIPLE and a THEORY OF COMPARISON-FACTORS

by

Werner C. Rheinboldt
Director and Research Professor
Computer Science Center
University of Maryland, College Park, Maryland

The work reported here was in part supported by the National Aeronautics and Space Administration under grant NsG-398. Acknowledgment is also given to the U. S. Army Research Office which stimulated this work indirectly through research grant DA-31-124-G563.

CONTENTS

ABSTRACT

21778

Recently, G. Aumann introduced the idea of the comparison-factor in approximation problems. This idea is here generalized to assume the form of a general estimation principle for non-negatively homogeneous real functions on linear spaces. A number of results are then proved concerning the finiteness of the correspondingly generalized comparison-factor. In addition, transformations of the underlying spaces are investigated which preserve the finiteness of these factors. The results are then applied to the special case of finite dimensional linear spaces and to a particular case which provides a connection to earlier results of other authors on best error bounds for optimal linear approximations.

Author

On a General Estimation Principle and a Theory of Comparison Factors

by Werner C. Rheinboldt

I. INTRODUCTION

In a recent paper [1], G. Aumann introduced a concept called the comparison-factor for linear approximation problems. The essential point behind Aumann's observations can be presented in the following form:

Let X be a real or complex linear space and q_1, q_2 two seminorms on X. Define a linear approximation problem by considering the null space $N(q_2) = \{x\in X \mid q_2(x) = o\}$ of q_2 as the subspace of approximating elements. Then, for given $x_o\in X$ an element $y_o\in N(q_2)$ has to be found such that

(1) $$q_1(x_o-y_o) = \inf_{y\in N(q_2)} q_1(x_o-y) = \delta(x_o)$$

The question about the existence and uniqueness of y_o shall not be of concern here; instead, upper bounds are to be determined for the best approximation error $\delta(x_o)$. Of course, without any further knowledge about x_o or q_1, q_2 etc., very little can be said about this problem, except that

$$\delta(x_o) \leqq q_1(x_o-y)$$

for any $y\in N(q_2)$. Suppose, therefore, that an additional estimate

(2) $$p(x_0) \leqq c$$

is known, where p is another seminorm on X with the property that $N(p)\subset N(q_2)$.

Then it is readily seen that

(3) $$\delta(x_o) \leqq \gamma c$$

where

(4) $$\gamma = \sup_{p(x)\leqq 1} \left[\inf_{q_2(y)=o} q_1(x-y) \right]$$

(For a generalized estimate of this type see Lemma 1 below.) The factor γ, which can, of course, be infinite, depends only on the three seminorms q_1, q_2, p, and not on x_o. In particular, when $q_2 \equiv p$, the additional condition $N(p) \subset N(q_2)$ is automatically satisfied and γ is only a function of the given approximation problem; in this case, G. Aumann calls γ the comparison-factor of this approximation problem.

In all cases, the significance of γ is the following: The estimate (2) is considered a known or "accessible" estimate for x_o and then the knowledge (and finiteness) of γ provides an estimate for the unknown best approximation error $\delta(x_o)$. Many well known bounds for approximation errors are examples for estimates of the type (3)/(4). For example, in [1] it is shown that the first theorem of D. Jackson in the theory of harmonic analysis has this form. For details we refer to [1].[1]

In line with these remarks, G. Aumann observes that a general study of the factor γ certainly appears to be an important problem of approximation theory. Such a study might include an investigation of the functional relationship between γ and the three seminorms q_1, q_2, p; another problem could be the

---

1] In [1] the seminorms q_1, q_2, p are also allowed to assume the value $+\infty$. For our purposes this generalization is of little importance, in particular, since all of the observations in [1] concern a subspace of X in which q_1, q_2 and p are finite.

development of methods for the computation or estimation of γ. In this general form the problem appears to have received little attention. However, a number of results about γ are obtained in the study of the related problem of the degree of approximation. In that case, a sequence $\{q_2^n\}$, $(n = 1,2,\ldots)$ of seminorms are considered such that $N(q_2^{n+1}) \supset N(q_2^n)$, and the behavior of $\delta_n(x_o)$ for $n \to \infty$ is investigated.

In some ways, G. Aumann's paper [1] can be regarded as the announcement of a "program" for the above stated general investigation of the factor γ. The present paper is intended to be a contribution to this program; at the same time, it was found necessary to free this entire problem from its setting of linear approximation problems. This leads in Section II to an extension of the estimation principle (3) and to a corresponding generalization of the definition of γ which in turn permits in Section III an attack on the important problem of the finiteness of γ. The results of Section III are based on the assumption that a certain set is bounded. In order to weaken this condition, transformations of the underlying spaces are considered in Section IV which leave the comparison-factors finite. In Section V the results of the earlier sections are applied to the special case of finite-dimensional spaces, and finally, Section VI connects some of the results of this paper with earlier work, notably by A. Sard [2] and H. Weinberger and W. Golomb [3].

Grateful acknowledgement is hereby given to the U. S. National Aeronautics and Space Administration which supported in part the preparation of this paper under grant NsG-398, and to the U. S. Army Research Office which stimulated this work indirectly through research grant D-31-124-G563. The author is also indebted to Professor J. Ortega for many stimulating discussions and for his critical and helpful comments.

II. THE LINEAR ESTIMATION PRINCIPLE

In the linear estimate (3) very little use is made of the fact that the quantity to be bounded is the best-approximation error $\delta(x_o)$. In fact, the crucial property needed in deriving (3) is the non-negative homogeneity of $\delta(x_o)$ and p. Furthermore, observe that instead of the condition $N(p) \subset N(q_2)$ we use only that $c = p(x_o) = 0$ implies $\delta(x_o) = 0$.

In order to free ourselves from the setting of linear approximation problems, we observe that the estimate (3) is a special case of a linear estimation principle which uses only the above properties of $\delta(x_o)$ and p.

Before phrasing this principle in the form of Lemma 1 below, we introduce the following notation:

Throughout this paper, X,Y,Z, etc. shall always denote real or complex linear spaces. For any subset $Q \subset X$, we denote by $R(Q)$ the real, linear space of all real-valued functions p defined on Q. Consider, in particular, a subset $C \subset X$ which is a cone (with vertex zero)[2], i.e., assume that $tC \subset C$ for all non-negative real t. Then for fixed given $\mu > 0$, $H_\mu(C)$ shall be the subset of all non-negative functions $p \in R(C)$ which are non-negatively homogeneous of degree μ, i.e.,

$$H_\mu(C) = \left\{p \in R(C) \,\middle|\, p(x) \geqq 0 \text{ and } p(tx) = t^\mu p(x) \text{ for any } t \geqq o \text{ and } x \in X\right\}$$

Evidently, $H_\mu(C)$ is a convex cone in $R(C)$. For the functions $p \in H_\mu(C)$ we introduce the sets

2] In the following, all cones shall always be assumed to have vertex zero, and to contain elements different from zero.

$$N(p) = \{ x \in C \mid p(x) = 0 \} ; \quad S(p) = \{ x \in C \mid p(x) \leqq 1 \}$$

and

$$\Gamma(p) = \{ x \in C \mid p(x) = 1 \} ; \quad S^o(p) = \{ x \in C \mid x = o \text{ or } o < p(x) \leqq 1 \}$$

Clearly $N(p)$ is a cone in C, $S(p) = S^o(p) \cup N(p)$ and $\lambda S(p) \subset S(p)$, $\lambda S^o(p) \subset S^o(p)$ for $0 \leqq \lambda \leqq 1$. $S(p)$ is radial in C, i.e., for each $x \in C$ there exists a $\delta = \delta(x) > o$ such that $\lambda x \in S(p)$ for $o \leqq \lambda \leqq \delta$. In fact, define $\delta = 1$ for $x \in N(p)$ and $\delta = p^{-1/\mu}$ if $p = p(x) > o$. Finally, it is easily checked that

$$S^o(p) = \bigcup_{o < \lambda \leqq 1} \lambda \Gamma(p)$$

Lemma 1: (The Linear Estimation Principle) Given a cone C in the linear space X and functions $r, p \in H_\mu(C)$, then

(5) $\quad \gamma_1 r(x) \leqq p(x) \quad (x \in C)$

where

$$\gamma_L = \inf_{y \in \Gamma(r)} p(y) \geqq 0 \quad , \text{ if } r,p \neq 0$$

and $\gamma_L = 0$ otherwise. Moreover, $\gamma_L > 0$ only if $N(p) \subset N(r)$. If this latter condition is satisfied, then

(6) $\quad r(x) \leqq \gamma_U p(x) \quad$ for $x \in C$

where

(7) $$\gamma_U = \sup_{y \in \Gamma(p)} r(y) \equiv \sup_{y \in S(p)} r(y) \text{ , and } 0 < \gamma_U \leqq \infty$$

if $r,p \neq 0$ and $\gamma_U = 0$ otherwise. If $r,p \neq 0$, then either $\gamma_U = \infty$ and $\gamma_L = o$, or γ_U is finite and $\gamma_L \gamma_U = 1$.
The factor $\gamma_U = \gamma_p(r,C)$ shall be called the comparison-factor of r with respect to p on the cone C.

Proof: The case $r,p \equiv 0$ is trivial. Assume $r,p \not\equiv 0$; then there exists an $x \in C$ such that $r_o = r(x) > 0$. But then $r(r_o^{-1/\mu} x) = 1$, and hence γ_L is well-defined and $\frac{1}{r_o} p(x) = p(r_o^{-1/\mu} x) \geqq \inf_{y \in \Gamma(p)} p(y)$.

This proves (5) for $x \notin N(r)$ and therefore for all $x \in C$ since the inequality clearly remains valid if $x \in N(r)$. If $x \in N(p)$, then $\gamma_L r(x) = 0$ and hence either $\gamma_L = 0$ or $x \in N(r)$. If the condition $N(p) \subset N(r)$ is satisfied, and $x \in C$ an element such that $p_o = p(x) > o$, then it follows similarly that γ_U is defined and that

$$\frac{1}{p_o} r(x) = r(p_o^{-1/\mu} x) \leqq \sup_{y \in \Gamma(p)} r(y) = \gamma_U$$

If $x \in N(p)$ then $x \in N(r)$ and hence (6) is valid for all $x \in C$. Moreover, $\gamma_U > o$ since otherwise $r(x) \equiv 0$. Clearly, $0 < \gamma_U = \sup_{y \in \Gamma(p)} r(y) \leqq$ $\leqq \sup_{y \in S(p)} r(y)$. For $y \in S^o(p)$ we now find that

$$r(y) = p_o r(p_o^{-1/\mu} y) \leqq p_o \gamma_U \leqq \gamma_U$$

and this inequality again remains valid for $y \in N(p)$, since then $r(y) = 0$. Therefore, $\sup_{y \in S(p)} r(y) \leqq \gamma_U$ and (7) is proven. Finally, from (5) it follows that $\gamma_L \gamma_U \leqq 1$ and hence $\gamma_U < \infty$ if $\gamma_L > o$. Similarly, (6) implies that $1 \leqq \gamma_L \gamma_U$ or $\gamma_L > o$ if $\gamma_U < \infty$. Together therefore, either $\gamma_U = \infty$ and $\gamma_L = o$ or $\gamma_U < \infty$ and $\gamma_L \gamma_U = 1$.

If instead of $N(p) \subset N(r)$ we even have $N(p) = N(r)$, then Lemma 1 states that

$$\gamma_o p(x) \leqq r(x) \leqq \gamma_1 p(x) \qquad (x \in C)$$

where $\gamma_1 = \gamma_p(r,C)$ and $\gamma_o = 1/\gamma_r(p,C)$ if $\gamma_r(p,C) < \infty$.

The relation between γ_L and γ_U in Lemma 1 permits us to concentrate the investigation on one of these factors. In this presentation we consider mainly the "upper" comparison-factor $\gamma_p(r,C)$. Lemma 1 evidently provides a non-trivial result only if we know that $\gamma_p(r,C)$ is finite. Hence, we shall be concerned primarily with the investigation of conditions for the finiteness of $\gamma_p(r,C)$. In the case $N(p) = N(r)$, such conditions will then also provide results about a non-trivial lower bound for r in terms of p.

<u>Examples:</u> (1a) As in Section I, let q_1, q_2, and p be three semi-norms on the linear space X and set

$$r(x) = \inf_{q_2(y)=o} q_1(x-y) = \delta(x)$$

Then p, $r \in H_1(X)$ and clearly $N(q_2) \subset N(r)$ since choosing an element x_o from the subspace $N(q_2)$ of approximating elements implies, of course, that the best approximation error $\delta(x_o)$ vanishes. Hence, if we assume that $N(p) \subset N(q_2)$, then

$$\delta(x) \leqq \gamma p(x)$$

where

$$\gamma = \sup_{p(x)\leqq 1} \left[\inf_{q_2(y)=0} q_1(x-y) \right]$$

and this is exactly the estimate (3)/(4).

(1b) Let X be an inner product space and A a linear operator with domain X and range in X. Set

$$p(x) = (x,x) \ , \ r(x) = |(Ax,x)|$$

Then clearly $p, r \in H_2(X)$ and $N(p) \subset N(r)$; hence

$$|(Ax,x)| \leqq \gamma(x,x) \qquad x \in X,$$

where

$$\gamma = \sup_{(x,x) \leqq 1} |(Ax,x)|$$

is finite if A is continuous on X.

(1c) Let X and Y be normed linear spaces and $A:X\to Y, B:X\to Y$ linear operators with domain X. Set

$$p(x) = \|Ax\| \quad , \quad r(x) = \|Bx\|$$

Then $p,r\in H_1(X)$. $N(p) \subset N(r)$ is equivalent to the assumption that $Ax = 0$ implies $Bx = 0$. In that case we have

$$\|Bx\| \leqq \gamma \, \|Ax\| \quad , \quad \gamma = \sup_{\|Ay\|\leqq 1} \|By\|$$

where, as usual, γ can be infinite. If $X \equiv Y$, A the identity, and B bounded, then γ is finite and equal to the norm of B.

(1d) Let X be the space of all complex $n \times n$ matrices and

$$r(x) = \max_{j,k} |x_{jk}|, \qquad p(x) = \max_{j} |\lambda_j|$$

where x_{jk} are the elements and $\lambda_1, \lambda_2, \ldots, \lambda_n$ the eigenvalues of the matrix $x\in X$. Then $r,p\in H_1(X)$ and $N(r) \subset N(p)$. Hence,

$$\max_{j} |\lambda_j| \leqq \gamma_1 \max_{j,k} |x_{jk}|$$

Let C be the cone of normal matrices in X. Then also $N(p) \subset N(r)$, since $p(x) = 0$, $x\in C$ implies the existence of a unitary matrix $u\in C$ such that $u^*x\,u = 0$, and therefore $r(x) = 0$. Thus

$$\max_{j,k} |x_{jk}| \leqq \gamma_2 \max_{j} |\lambda_j| \qquad x\in C$$

It is well known that $\gamma_1 = n$, and $\gamma_2 = 1$.

We end this section with a brief note about the set $H_\mu(C)$. The definitions

$$r \prec p \quad \text{if and only if} \quad r,p\in H_\mu(C) \text{ and } N(p) \subset N(r)$$

and

$$r \ll p \quad \text{if and only if} \quad r \prec p \text{ and } \gamma_p(r,C) \text{ is finite}$$

introduce partial orderings in $H_\mu(C)$. In fact, it is readily seen that both relations are reflexive and transitive. Moreover, with $r,p\in H_\mu(C)$ also $\min(r,p)$ and $\max(r,p)$ are functions of $H_\mu(C)$ and evidently, under both orderings, $\max(r,p)$ is a least upper bound of r,p and $\min(r,p)$ a greatest lower bound. Yet these partial orderings are not fully compatible with the linear structure of $H_\mu(C)$. Clearly, $r \prec p$ implies that $\alpha r \prec \alpha p$ for any real $\alpha \geq 0$, and $r+q \prec p+q$ for any $q \in H_\mu(C)$, and the same holds true for the stronger ordering $\prec\!\prec$. But it can be shown with easy examples that $r \prec p$ and q, $r-q$, $p-q \in H_\mu(C)$ does not generally imply that $r-q \prec p-q$. For the questions here under consideration, this fact appears to limit the usefulness of the above orderings.

III. BASIC FINITENESS THEOREMS FOR COMPARISON-FACTORS.

As we mentioned before, Lemma 1 has no particularly significant value unless we know that the comparison-factor $\gamma_p(r,C)$ is finite. The following lemma casts this problem into a somewhat different form:

Lemma 2: Given a cone C in the linear space X and functions $r,p \in H_\mu(C)$ such that $N(p) \subset N(r)$. Then $\gamma_p(r,C)$ is finite if and only if $S(r)$ absorbs $\Gamma(p)$, i.e., if and only if there exists a positive constant k such that $\Gamma(p) \subset kS(r)$.

Proof: If $\gamma = \gamma_p(r,C)$ is finite, then (6) implies immediately that $\Gamma(p) \subset kS(r)$ for any $k > o$ with $k^\mu \geq \mu$. Reversely, if $S(r)$ absorbs $\Gamma(p)$, then it follows for $x \in \Gamma(p)$ that $x=ky, y \in S(r)$ and hence $r(x) \leq k^\mu$ and therefore $\gamma \leq k^\mu$.

The condition that $S(r)$ absorb $\Gamma(p)$ appears to be weaker than it actually is. This is shown by the following:

Lemma 3: Let Q be a subset of X and $\alpha > o$ and $\beta > o$ any constants; then the following three conditions are equivalent:

(i) $\beta S(r) \supset Q \cap \alpha\Gamma(p)$

(ii) $\beta S(r) \supset Q \cap \alpha S^o(p)$

(iii) $\beta S(r) \supset Q \cap \alpha S(p)$

The proof is immediate. From (i) it follows that

$\alpha\lambda\,\Gamma(p)\cap Q\subset\beta S(r)$, $o<\lambda\leqq 1$, and hence that $\alpha S^{o}(p)\cap Q = \alpha\Big(\bigcup_{o<\lambda\leqq 1}\lambda\,\Gamma(p)\Big)\cap Q\subset\beta S(r)$.

(ii) implies that

$$Q\cap\alpha S(p) = Q\cap\Big(\alpha S^{o}(p)\cup N(p)\Big)\subset\Big(Q\cap\alpha S^{o}(p)\Big)\cup\Big(Q\cap N(p)\Big)$$

$$\subset\beta S(r)\cup N(p)\subset\beta S(r)\cup N(r) = \beta S(r)$$

and it is clear that (i) follows from (iii)

Accordingly, without loss of generality, we could have phrased Lemma 2 in the form: $\gamma_p(r,C)$ is finite if and only if $S(r)$ absorbs $S(p)$.

Obviously, general results about the finiteness of γ can be expected only if additional conditions are placed on X as well as on the functions r,p. As a first and basic result we obtain the following:

<u>Theorem 1:</u> Consider a cone C in the topological linear space X and functions $r,p\subset H_\mu(C)$ with $N(p)\subset N(r)$ such that $\Gamma(p)$ is a bounded set. Then $\gamma_p(r,C)$ is finite if for some $k\geqq 1$ the origin is in the relative interior of $kS(r)$ with respect to $S(p)$. Hence, $\gamma_p(r,C)$ is clearly finite if r is continuous at zero in the relative topology on $S(p)$, or even on C.

<u>Proof:</u> By assumption there exists a neighborhood U of zero such that

(8) $\quad kS(r)\supset U\cap S(p)$

By definition, U absorbs the bounded set $\Gamma(p)$, i.e., $c\,\Gamma(p)\subset U$ for some c with $0<c\leqq 1$. Hence $kS(r)\supset U\cap S(p)\supset c\,\Gamma(p)\cap S(p)=c\,\Gamma(p)$, or $S(r)$ absorbs $\Gamma(p)$ and by Lemma 2, $\gamma_p(r,C)$ is finite.

<u>Note:</u> Because of Lemma 3, the basic assumption (8) of this theorem is equivalent to $kS(r)\supset U\cap S^{o}(p)$ as well as $kS(r)\supset U\cap\Gamma(p)$. Moreover, boundedness of $\Gamma(p)$ is equivalent to boundedness of $S^{o}(p)$. In fact, since $\Gamma(p)\subset S^{o}(p)$ we need to show only that $S^{o}(p)$ is bounded

if $\Gamma(p)$ is bounded. In that case, let W be any neighborhood of zero and $V \subset W$ any balanced neighborhood of zero, then $c\,\Gamma(p) \subset V$ for some $c>0$, and hence $c\lambda\,\Gamma(p) \subset \lambda V \subset W$ for $o<\lambda \leq 1$ and finally $cS^{o}(p) \subset W$.

In Theorem 1 a non-topological conclusion - namely the finiteness of γ - is deduced from a set of topological premises. From a theoretically strict viewpoint this may be somewhat unsatisfactory. But in many applications it does appear to be a very natural setting for this problem to assume that the underlying space X is a topological linear space. Moreover, the theorem can also be rephrased to state that γ is finite if there exists <u>any</u> vector topology on X such that $\Gamma(p)$ is bounded, and zero is in the interior of $kS(r)$ relative to $S(p)$. The question then arises whether there are some "natural" topologies connected with p and r for which the Theorem assumes a straightforward set-theoretic form. As we shall se later, there are indeed special cases of p when such a natural topology exists, but the corresponding result for γ is then essentially equivalent to that of Lemma 2.

For the application of Theorem 1 it is often desirable to replace the conditions on r and p by stronger but more "usable" assumptions.

A very natural condition for a function $p \in H_{\mu}(C)$ is that of the convexity of $S(p)$. If the cone $C \subset X$ is convex, i.e., if $C+C \subset C$, and if for $\mu \geq 1$ the function $p \in H_{\mu}(C)$ is subadditive, then $S(p)$ and $N(p)$ are clearly convex sets. More generally, we can give the following necessary and sufficient condition: If $C \subset X$ is a convex cone and $p \in H_{\mu}(C)$ $(\mu>0)$, then $S(p)$ is convex if and only if $p\big(\lambda x+(1-\lambda)y\big) \leq \max\big(p(x), p(y)\big)$ for every $x,y \in C$ and

$0 \leq \lambda \leq 1$. This result will be contained in the statement of Lemma 4 below. See also W. Fenchel [4], p. 117.

Convexity is a fairly strong condition for S(p) and in many instances a weaker condition will suffice. In general, a set $Q \subset X$ shall be called quasiconvex of degree α if there exists a constant $\alpha \geqq 1$ such that $\lambda Q + (1-\lambda)Q \subset Q$ for all $0 \leq \lambda \leq 1$. Clearly, a set Q is convex if and only if it is quasiconvex of degree 1; and, if Q is a cone, quasiconvexity of Q implies that Q is convex.

Let $C \subset X$ be a convex cone, then $p \in H_\mu(C)$ is called a quasiconvex function of degree α if the set S(p) is quasiconvex of degree α.

<u>Lemma 4:</u> Let $C \subset X$ be a convex cone, then $p \in H_\mu(C)$ $(\mu > 0)$ is quasiconvex of degree α if and only if

(9) $$p(\lambda x + (1-\lambda)y) \leq \alpha^\mu \max (p(x), p(y))$$

for $x, y \in C$ and $0 \leq \lambda \leq 1$. If p is quasiconvex, then N(p) is convex.

<u>Proof:</u> If $p \in H_\mu(c)$ satisfies (9), then $x, y \in S(p)$ implies that $p(\lambda x + (1-\lambda)y) \leq \alpha^\mu$ for $0 \leq \lambda \leq 1$, or $\lambda x + (1-\lambda)y \in \alpha S(p)$, i.e., S(p) is quasiconvex of degree α. Reversely, let S(p) be quasiconvex of degree α. Then clearly

(10) $$N(p) + N(p) \subset S(p) = S^0(p) \cup N(p)$$

If there exist $x, y \in N(p)$ such that $o < p_o = p(x+y) \leq 1$, then also $tx, ty \in N(p)$ for $t > 0$, but $tx + ty \notin S(p)$ for $t^\mu > (1/p_o)$. This contradicts (10) and hence $N(p) + N(p) \subset N(p)$, or N(p) is convex. For given $x, y \in C$, now set $\tau = \max(p(x), p(y))$. If $\tau = 0$, then $x, y \in N(p)$ and (9) holds, since the convexity of N(p) implies that $p(\lambda x + (1-\lambda)y) = 0$ for $0 \leq \lambda \leq 1$. Let therefore $\tau > 0$; then $x, y \in \tau^{1/\mu} S(p)$, and since $\tau^{1/\mu} S(p)$ is again quasiconvex of degree α, it follows that $\lambda x + (1-\lambda)y \in \alpha \tau^{1/\mu} S(p)$ or $p(\lambda x + (1-\lambda)y) \leq \alpha^\mu$, which proves (9).

Let X be real <u>or</u> complex and $C \subset X$ a convex cone. If $C = -C$ then C is a <u>real</u> linear space and in that case $p \in H_\mu(C)$ shall be called real-symmetric if $p(x) = p(-x)$ for all $x \in C$. If X is complex and $x \in C$ implies $\lambda x \in C$ for all $|\lambda| = 1$, then C is a (complex) linear subspace of X and $p \in H_\mu(C)$ shall be called complex symmetric if $p(\lambda x) = p(x)$ for all $x \in C$ and $|\lambda| = 1$.

Consider the case when $C = -C$ and $p \in H_\mu(C)$ $(p \neq 0)$ is quasiconvex and real symmetric. Then $S(p)$ is balanced in the real linear space C, i.e., $\lambda S(p) \subset S(p)$ for $-1 \leq \lambda \leq +1$. Furthermore, the family of sets

$$\mathfrak{U} = \{U_t \mid U_t = tS(p),\ t > 0\}$$

is a local base of zero for a vector topology τ_p on C. In fact, the sets U_t are balanced (in the real space C) and absorb every point of C. Moreover $U_t \subset U_{t_1} \cap U_{t_2}$ for $t \leq \min(t_1, t_2)$ and $U_t + U_t \subset U_{t_0}$ for $t \leq t_0/2\alpha$. This real topological linear space C is certainly locally bounded. Moreover, this topology τ_p of C is equivalent to the one induced by using as local base of zero only the sets U_t with rational $t > 0$. Hence, τ_p has a countable local base at zero and is therefore semimetrizable. In other words, there exists an invariant semimetric such that each sphere around zero is balanced and that this semimetric induces the topology τ_p on C. If $\alpha = 1$, then τ_p is a locally convex topology and is equivalent to the seminorm-topology induced in C by the Minkowski functional of $S(p)$.

Of course, similar results hold in the case that X is complex and if $\lambda C = C$ for $|\lambda| = 1$ and p is quasiconvex and complex-symmetric.

Examples: (2a) For fixed $\mu > 0$, let X be the space of all real sequences $x = \{\xi_n, n = 1, 2, \ldots\}$ such that $\sum_{n=1}^{\infty} |\xi_n|^\mu$ converges. Then $p(x) = \left(\sum_{n=1}^{\infty} |\xi_n|^\mu\right)^{1/\mu} \in H_1(X)$. For $0 < \mu < 1$, p is quasiconvex of degree

$2^{(1-\mu)/\mu}$ while for $\mu \geqq 1$, $S(p)$ is, of course, convex.

(2b) Let X be the two-dimensional real number space. For $x=(\xi_1, \xi_2) \in X$ define $p(x)=2(|\xi_1|+|\xi_2|)$ if $\operatorname{sgn}\xi_1=\operatorname{sgn}\xi_2$ and $p(x)=|\xi_1|+|\xi_2|$ otherwise. Then p is quasiconvex of degree 2.

(2c) The last example is a special case of the following general result: Let X be a topological linear space and assume $p \in H_\mu(X)$ has the property that $S(p)$ is a bounded set which contains zero as interior point. Then p is quasiconvex of some degree $\alpha \geqq 1$. In fact, these assumptions assure that there is a neighborhood U of zero such that $S(p) \supset U$. Now a balanced neighborhood V of zero can be chosen with $V+V \subset U$, and since $S(p)$ is bounded, there exists a constant $\alpha \geqq 1$ such that $S(p) \subset \alpha V$. But then $\lambda S(p)+(1-\lambda) S(p) \subset \alpha V+\alpha V \subset \alpha U \subset \alpha S(p)$ for $o \leqq \lambda \leqq 1$, i.e., p is quasiconvex of degree α.

We return to our question of the finiteness of γ. For a linear space X, a quasiconvex and symmetric function $p \in H_\mu(X)$ induces a "natural" topology τ_p on X. Since the set $S(p)$ is bounded under τ_p, we accordingly arrive at a very simple specialization of Theorem 1. As was indicated earlier, however, a closer inspection quickly shows that this special case of Theorem 1 is in fact only a differently phrased version of Lemma 2. Nevertheless, the following Corollary plays an important role in many applications:

<u>Corollary 1.1</u>: Given a linear space X and a function $p \in H_\mu(X)\,(p \not\equiv 0)$ which is (real or complex) symmetric and quasiconvex. Then for any $r \in H_\mu(X)$ with $N(p) \subset N(r)$, $\gamma_p(r,X)$ is finite if and only if r is continuous at the origin under the topology τ_p induced by p on X.

Note that the topology τ_p may be rather unusual if X is complex and p only real symmetric. In that case, τ_p is a "real" vector topology on the space X, also considered as a real linear space.

Corollary 1.1. of course covers the case described in Example (1b). In fact, in that case γ will be finite if the linear operator A is continuous, since $r(x)=(Ax,x)$ is then continuous under the norm topology, here the τ_p topology.

For the application of Theorem 1 and Corollary 1.1 it is often useful to know when a function $r\in H_\mu(X)$ is continuous at the origin. For quasiconvex functions the following result holds:

Theorem 2: Let X be a topological linear space and $r\in H_\mu(X)$ quasiconvex. Then r is continuous at zero if either one of the following two conditions is satisfied:

(a) The interior of $S(r)$ is not empty

(b) X is a complete semimetrizable space and r is lower semicontinuous on X[3]

Proof: (a) By assumption there exists a neighborhood U of zero and a point $z\in S(r)$ such that $S(r)\supset z+U$. Since $-z\in kS(r)$ for $k^\mu \geqq r(-z)$ we have $U\subset(z+U)+(-z)\subset S(r)+kS(r)\subset k_1S(r)$ for $k_1=2\alpha\max(1,k)$. Hence r is continuous at zero.

(b) The Baire category theorem applies to X and implies that X is of second category in itself. Since $X=\bigcup_{n=1}^{\infty} nS(r)$, it follows that for some integer n_o, $n_oS(r)$ is of second category in X, hence $S(r)$ must also be of second category in X. The lower

3] Recall that for a topological linear space X, a function $r\in R(X)$ is called lower semicontinuous on X if for each real t the set $\left\{x\in X \mid r(x)\leqq t\right\}$ is closed. Hence, clearly, $r\in H_\mu(X)$ is lower semicontinuous on X if and only if $S(r)$ is closed.

semicontinuity of r implies that $S(r)$ is closed. But then the interior of $S(r)$ cannot be empty since otherwise $S(r)$ would consist only of boundary points and would therefore be nowhere dense in X. Now Part (a) applies and the proof is complete.

<u>Note:</u> Condition (b) of this theorem could have been replaced by the following weaker condition:

(b') X is complete and semimetrizable and $S(r) \subset kS(r)$ for some $k \geq 1$. In fact, the same proof shows that $S(r)$ is of second category in X and has, therefore, non-empty interior. Hence, $kS(r)$ and also $S(r)$ have non-empty interior.

Theorem 2, Part (a) and Theorem 1 together yield the following result: If $r,p \in H_\mu(X)$, $r,p \not\equiv 0$ and $N(p) \subset N(r)$, then $\gamma_p(r,X)$ is finite if $\Gamma(p)$ is bounded and if there exists a point $z \notin S(r)$ such that $kS(r) \supset (z+U)$ for some $k \geq 1$. It is natural to ask whether γ remains finite if we only know that z is interior to $kS(r)$ relative to $S(p)$. The following theorem gives a partial answer:

<u>Theorem 3:</u> For the topological linear space X let $r,p \in H_\mu(X)$ be such that $N(p) \subset N(r)$ and that both r and p are quasiconvex. Assume that $\Gamma(p)$ is bounded and that there exists a point $z \in S^0(p)$ which is interior to $kS(r)$, for some $k \geq 1$, in the relative topology on $\beta S(p)$ where β is the degree of quasiconvexity of p. Then $\gamma_p(r,X)$ is finite.

<u>Proof:</u> There exists a neighborhood U of zero such that $kS(r) \supset (z+U) \cap \beta S(p)$. Together with $\Gamma(p)$, also $S^0(p)$ and $S^0(p) - S^0(p)$ are bounded sets. Hence $c(S^0(p)-S^0(p)) \subset U$ for some c, $0 < c \leq 1$. Then

$$kS(r) \supset (z+U) \cap \beta S(p) \supset (z + c(S^0(p) - S^0(p))) \cap \beta S(p) \supset$$

$$\supset ((1-c)z + c S^0(p)) \cap \beta S(p) = (1-c)z + c S^0(p)$$

The last relation follows directly from the fact that p is quasiconvex of degree β. As before, we know that $-(1-c)z \in k_1 S(r)$ where we can choose $k_1 \geqq k$. Hence,

$$c\Gamma(p) \subset cS^{\circ}(p) \subset (1-c)z + cS^{\circ}(p) + (-(1-c)z) \subset kS(r) + k_1 S(r) \subset 2\alpha k_1 S(r)$$

or S(r) absorbs $\Gamma(p)$ and $\gamma_p(r,X)$ is finite.

If S(p) is not only quasiconvex but convex, then the main condition of this last theorem can be replaced by a category assumption for p. In fact, we find

Corollary 3.1: On the topological linear space X consider $p,r \in H_\mu(X)$ with $N(p) \subset N(r)$ such that p is quasiconvex of degree 1 and that $\Gamma(p)$ is bounded and of second category in itself. If r is quasiconvex of arbitrary degree and lower semicontinuous on X, then $\gamma_p(r,X)$ is finite.

Proof: Set $Q_n = nS(r) \cap \Gamma(p)$, then $\Gamma(p) = \bigcup_{n=1}^{\infty} Q_n$, and since $\Gamma(p)$ is of second category in itself, there exists a neighborhood U of zero and a point $z \in \Gamma(p)$ such that $\overline{Q}_{n_o} \supset (z+U) \cap \Gamma(p)$ and therefore $n_o S(r) = \overline{n_o S(r)} \supset \overline{Q}_{n_o} \supset (z+U) \cap \Gamma(p)$ and Theorem 3 applies.

Note: As in Theorem 2, the condition of lower semicontinuity of r could have been replaced by the assumption that $\overline{S(r)} \subset kS(r)$ for some $k \geqq 1$.

The result of Corollary 3.1 represents some variation of the so-called Absorption Theorem (see, e.g. J. Kelley, I. Namioka et al, [5], p. 90). In fact, in a similar manner we could have proved the following result: Let Q be a closed, quasiconvex set in a topological linear space X. Assume that Q absorbs each point of the set $P \cup (-P)$ where P is a bounded and convex set which is of second category in itself, then Q absorbs P.

IV. MAPPING THEOREMS

The basic assumption for all theorems of the last Section was the boundedness of $\Gamma(p)$. This is a fairly stringent condition which is not always satisfied even if γ is finite. Consider, for example the simple case when p is a seminorm on a finite dimensional normed linear space. If $N(p)$ contains elements different from zero, then $\Gamma(p)$ is clearly unbounded, while on the other hand we shall see later that in this case γ is already finite if r and p are both seminorms.

In order to find more general finiteness conditions for comparison-factors, transformations will now be considered which leave these factors finite. In general, let $\hat{C} \subset \hat{X}$ and $C \subset X$ be cones in the linear spaces $\hat{X}$ and X. Consider the sets $H_{\hat{\mu}}(\hat{C})$ and $H_{\mu}(C)$ for fixed $\hat{\mu} > 0$ and $\mu > 0$, and introduce on both the partial ordering " $\ll$ " defined at the end of Section 2. If then the mapping $F: H_{\hat{\mu}}(\hat{C}) \rightarrow H_{\mu}(C)$ is " $\ll$ "-isotone, i.e., if

$$\text{(11)} \qquad \hat{r} \ll \hat{p} \text{ for } \hat{r}, \hat{p} \in H_{\hat{\mu}}(\hat{C}) \text{ implies that } r = F\hat{r} \ll F\hat{p} = p$$

then $\gamma_p(r,C)$ is of course finite if $\gamma_{\hat{p}}(\hat{r},\hat{C})$ is finite. Therefore, the problem is to determine when such a mapping F is isotone under " $\ll$ ". Clearly, F must satisfy the following three necessary conditions: (i) F is non-negatively homogeneous of degree $\mu/\hat{\mu}$, (ii) for each $\hat{p} \in H_{\hat{\mu}}(\hat{C})$ it follows that $(F\hat{p})(x) \geq 0$ for all $x \in C$, (iii) if $\hat{r}, \hat{p} \in H_{\hat{\mu}}(\hat{C})$ and $N(\hat{p}) \subset N(\hat{r})$ then $N(F\hat{p}) \subset N(F\hat{r})$.

Instead of investigating general conditions under which a mapping F is " $\ll$ " -isotone, we shall concern ourselves here only with a special class of such mappings, namely those which are induced by certain transformations between C and $\hat{C}$. A more comprehensive investigation of general " $\ll$ " -isotone mappings is planned for a later paper.

<u>Theorem 4:</u> Given the linear spaces X and $\hat{X}$ and a cone $C \subset X$. Let $G: C \to \hat{X}$ be a non-negatively homogeneous mapping of degree $\omega > 0$, and consider the cone $GC = \hat{C} \subset \hat{X}$. Fix constants $a > 0$, $\nu > 0$, and $\hat{\mu} > 0$; for each $\hat{p} \in H_{\hat{\mu}}(\hat{C})$ the function $p(x) \equiv (F\hat{p})(x) = a\,\hat{p}^{\nu}(Gx)$ is then contained in $H_{\mu}(C)$ where $\mu = \hat{\mu}\nu\omega$. Furthermore, the mapping $F\hat{p} = p$ from $H_{\hat{\mu}}(\hat{C})$ into $H_{\mu}(C)$ is "$\prec\!\prec$"-isotone, and more specifically, if $\hat{r}, \hat{p} \in H_{\hat{\mu}}(\hat{C})$ and $N(\hat{p}) \subset N(\hat{r})$ then $r, p \in H_{\mu}(C)$, $N(p) \subset N(r)$ and either $\hat{\gamma} = \gamma_{\hat{p}}(\hat{r}, \hat{C})$ and $\gamma = \gamma_p(r,C)$ are both infinite or $\gamma = \hat{\gamma}^{\nu}$.

<u>Proof:</u> It is readily seen that $p = F\hat{p} \in H_{\mu}(C)$ if $\hat{p} \in H_{\hat{\mu}}(\hat{C})$. Assume therefore that $\hat{r}, \hat{p} \in H_{\hat{\mu}}(\hat{C})$ and $N(\hat{p}) \subset N(\hat{r})$. Then $x \in N(p)$ implies that $a\hat{p}^{\nu}(Gx) = p(x) = 0$ or $Gx \in N(\hat{p}) \subset N(\hat{r})$ and hence $r(x) = a\hat{r}^{\nu}(Gx) = 0$ or $x \in N(r)$. Since $\hat{C} = GC$ there exists for every $\hat{x} \in \hat{C}$ an $x \in C$ such that $\hat{x} = Gx$. Hence

$$\hat{r}(x) = \hat{r}(Gx) = \left(\tfrac{1}{a}\, r(x)\right)^{1/\nu} \le \gamma^{1/\nu}\left(\tfrac{1}{a}\, p(x)\right)^{1/\nu} = \gamma^{1/\nu}\,\hat{p}(Gx) = \gamma^{1/\nu}\,\hat{p}(x)$$

or $\hat{\gamma} \le \gamma^{1/\nu}$, and $\gamma = \infty$ if $\hat{\gamma} = \infty$. Reversely,

$$r(x) = a\,\hat{r}^{\nu}(Gx) = a\,\hat{r}^{\nu}(x) \le a\,\hat{\gamma}^{\nu}\,\hat{p}^{\nu}(\hat{x}) = \hat{\gamma}^{\nu}\,a\,\hat{p}^{\nu}(Gx) = \hat{\gamma}^{\nu}\,p(x)$$

implies that $\gamma \le \hat{\gamma}^{\nu}$ or $\hat{\gamma} = \infty$ if $\gamma = \infty$. Together therefore, either both $\hat{\gamma} = \infty$ and $\gamma = \infty$, or $\gamma = \hat{\gamma}^{\nu}$.

As a first simple consequence of this theorem we see that the degree μ of homogeneity has no influence on the problem of the finiteness of the comparison-factor. In particular, without loss of generality, it is always possible to consider instead of $p, r \in H_{\mu}(C)$ the functions $p^{1/\mu}, r^{1/\mu} \in H_1(C)$. However, for most practical applications it is usually more advantageous to work directly with the original functions in $H_{\mu}(C)$.

<u>Example (4a):</u> Let X be the space of all $n \times n$ complex matrices. The Hadamard determinant theorem states that

$$r(x) \equiv |\det(x)|^2 \le \prod_{j=1}^{n} \sum_{k=1}^{n} |x_{jk}|^2 \equiv p(x) \tag{12}$$

for all $x \in X$, i.e., that $\gamma_p(r,X) = 1$. Let C be the cone of all non-singular matrices. Since $r(x) = 0$ whenever $x \notin C$, we have to prove only that $\gamma_p(r,C) = 1$. If $x \in C$, then x^*x is positive-definite, hermitian, and can therefore be uniquely decomposed in the form $x^*x = z^*z$ where z is a non-singular, upper-triangular matrix with non-negative real elements on the main diagonal. Let $C_T(=\hat{C}) \subset X$ be the cone of all such matrices; then the mapping $G: x \in C \to z \in C_T$ is non-negatively homogeneous of first degree, and in this case it happens that $r(Gx) = r(X)$ and $p(Gx) = p(x)$ for all $x \in C$, i.e., that $\hat{p} \equiv p$ and $\hat{r} \equiv r$. In fact, introduce the auxiliary functions $p_o(x) = \prod_{j=1}^{n} |x_{jj}|$ and $r_o(x) = |\det(x)|$, then for all $x \in C$

$$r(x) = (r_o(x))^2 = r_o(x^*x) = r_o(z^*z) = r(z) = r(Gx)$$

$$p(x) = p_o(x^*x) = p_o(z^*z) = p(z) = p(Gx)$$

For $z \in C_T$ we now have

$$p(z) = \prod_{j=1}^{n} \sum_{k=1}^{n} |z_{jk}|^2 \geq \prod_{j=1}^{n} |z_{jj}|^2 = (\det(z))^2 = r(z)$$

with equality holding if $z \in C$ is diagonal. Therefore $\gamma_p(r,C_T) = 1$ and hence by Theorem 4, $\gamma_p(r,C) = 1$. Note that in this example, Theorem 4 is actually applied twice. From $\gamma_p(r,C)$ it is first deduced that the comparison-factor of r_o with respect to p_o on the cone of all positive definite, hermitian matrices has the value one. Using the mapping $x \to x^*x$, a second application of Theorem 4 then yields the final result.

(4b) Let X be a Hilbertspace and $A: X \to S$ a linear, bounded, self-adjoint and positive definite operator such that $m(x,x) \leqq (Ax,x) \leqq M(x,x)$ for all $x \in X$ with $M > m > 0$. Set $r(x) = (Ax,x)(A^{-1}x,x)$, $p(x) = (x,x)^2$, then the Kantorovich inequality states that $r(x) \leqq [(M+m)^2/4Mm]p(x)$, for $x \in X$. Let X be the two-dimensional (complex) number space and consider the composite mapping

$$G: \quad x \in X \longrightarrow z = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \in X \quad \text{where} \quad \xi^2 = \alpha(Ax,x) + \beta(A^{-1}x,x)$$
$$\text{and} \quad \eta^2 = \gamma(Ax,x) + \delta(A^{-1}x,x)$$

from X into $\hat{X}$ where $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is some non-singular, real 2×2 matrix with the property that $U \begin{pmatrix} m \\ 1/M \end{pmatrix}$ is a positive vector, i.e., that $\xi > 0$, $\eta > 0$. Then the linear operators $K^2x = \alpha Ax + \beta A^{-1}x$, and $L^2x = \gamma Ax + \delta A^{-1}x$ are bounded, self-adjoint and positive definite on H, and hence K and L exist and have the same properties. Using these operators K and L, the mapping G can be represented in the form $Gx = \begin{pmatrix} \|Kx\| \\ \|Lx\| \end{pmatrix}$.

For $z = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \in \hat{X}$ consider the functions

$\hat{r}(z) = (Bz,z)(B^{-1}z,z)$ where $B = \begin{pmatrix} M & 0 \\ 0 & m \end{pmatrix}$ and $\hat{p}(z) = \left(\xi^2 + \eta^2\right)^2$; then

$$\hat{r}(z) = \left(M\xi^2 + m\eta^2\right)\left(M^{-1}\xi^2 + m^{-1}\eta^2\right) = \hat{p}(z) + \frac{(M-m)^2}{Mm}\xi^2\eta^2 \leq$$
$$\leq \hat{p}(z)\left(1 + \frac{(M-m)^2}{4Mm}\right) = \frac{(M+m)^2}{4Mm}\hat{p}(z)$$

with equality holding if and only if $\xi^2 = \eta^2$. Hence $\gamma_{\hat{p}}(\hat{r},\hat{X}) = \frac{(M+m)^2}{4Mm}$. Now $\hat{r}(Gx) = (M K^2x + m L^2x, x)\left(\frac{K^2x}{M} + \frac{L^2x}{m}, x\right)$ and hence for $U^{-1} = \begin{pmatrix} M & m \\ M^{-1} & m^{-1} \end{pmatrix}$ we obtain $\hat{r}(Gx) = r(x)$. It is quickly checked that this choice of U is permissible. Introduce $p_o(x) = (\|Kx\|^2 + \|Lx\|^2) = \hat{p}(Gx)$, then Theorem 4 yields

$$\gamma_{p_o}(r,X) = \frac{(M+m)^2}{4Mm} \quad \text{or}$$

$$(Ax,x)(A^{-1}x,x) \leq \frac{(M+m)^2}{4Mm}\left(\|Kx\|^2 + \|Lx\|^2\right) \quad (x \in X)$$

This is a generalized form of an inequality of J. Diaz and F. Metcalf [6]. A simple calculation shows that

$$(x,x) = \|Kx\|^2 + \|Lx\|^2 + \frac{M}{m}(Kx,Lx) + \frac{m}{M}(Lx,Kx)$$

which proves the Kantorovich inequality. At the same time it follows that in the Kantorovich inequality, equality holds if and only if $\|Kx\| = \|Lx\|$ and $(Kx, Lx) = 0$. For finite dimensional X this was shown by P. Henrici [7].

These two examples are not meant to give a "new" proof for the two inequalities, but are rather intended to show that these inequalities represent instances of the general theory here under discussion. The type of reasoning used in Example (4a) appears to be quite generally applicable. In fact, upon examining a number of the known matrix inequalities it turned out that in each case some suitable non-negatively homogeneous matrix transformation existed such that on the range-cone the particular inequality assumed a trivial form and that Theorem 4 assured the validity in the general case.

The mapping $F: H_{\hat{\mathcal{C}}}(\hat{C}) \longrightarrow H_{\mathcal{C}}(C)$ considered in Theorem 4 is evidently one-to-one. If we denote the range of F by $\mathcal{R}(F)$, there exists an inverse mapping $F^{-1}: \mathcal{R}(F) \longrightarrow H_{\hat{\mathcal{C}}}(\hat{C})$ and F^{-1} is again "$\ll$"-isotone. In fact, if $p, r \in \mathcal{R}(F)$, $N(p) \subset N(r)$ and $\gamma = \gamma_p(r, C)$ is finite, it follows from the proof of Theorem 4 that

$$\hat{r}(\hat{x}) \equiv (F^{-1}r)(\hat{x}) \leq \gamma^{1/\nu}(F^{-1}p)(\hat{x}) \equiv \gamma^{1/\nu} p(x) \text{ for all } \hat{x} \in \hat{C}.$$

Hence $N(\hat{p}) \subset N(\hat{r})$ and clearly $\hat{\gamma} = \gamma_{\hat{p}}(\hat{r}, \hat{C})$ is finite.

This observation will be really useful only if it is possible to say what functions of $H_{\hat{\mathcal{C}}}(\hat{C})$ are contained in $\mathcal{R}(F)$. In this paper we shall not consider this question in some generality, but instead we will only present a special case, which is important for the applications. For this purpose the following concept will be needed:

Let $Q \subset X$ be a subset of the linear space X, then $p \in R(Q)$ is said to satisfy a -condition on Q if there exists a constant $\alpha \geq 1$ such that

$$(13) \qquad |p(x) - p(y)| \leq \alpha p(x-y) \quad \text{whenever } x, y, x-y \in Q$$

A function $p \in R(Q)$ which satisfies (13) has the following properties:

(1) If $0 \in Q$, then $p(x) \geq 0$ for all $x \in Q$, and if $p(0) = 0$, then $\alpha p(x) \geq p(-x)$ whenever $x, -x \in Q$

(2) If Q is a convex cone and $p \in H_{\mu}(Q)$, then p is quasi-convex of degree $\beta = \alpha^{1/\mu}$, and if $\alpha = 1$, then p is a seminorm on Q. Reversely, every seminorm of X satisfies a U-condition with $\alpha = 1$.

(3) If X is a topological linear space and Q a linear subspace, then the continuity of p at zero implies uniform continuity of p on all of Q. This fact prompted the use of the letter U for condition (13).

The proof of these three properties is immediate and has been omitted here. Using this concept we obtain:

Lemma 5: Given the linear spaces X and $\hat{X}$ and a linear mapping L from X onto $\hat{X}$. Similar to Theorem 4 consider the mapping $p(x) \equiv (F\hat{p})(x) = a\,\hat{p}^{\nu}(Lx)$ from $H_{\hat{\mu}}(\hat{X})$ into $H_{\mu}(X)$ where $\mu = \hat{\mu}\nu$. Then the range of F contains the functions $p \in H_{\mu}(X)$ which satisfy a U-condition on X and have the property that $N(L) \subset N(p)$ where $N(L) = \left\{ x \in X \mid Lx = 0 \right\}$.

The proof is immediate. In fact, let p satisfy the two conditions of the lemma, then $\hat{p}(\hat{x}) \equiv \frac{1}{a}\, p^{1/\nu}(x)$, where $\hat{x} = Lx$, is a function of $H_{\hat{\mu}}(\hat{X})$. To see this we need to show only that $\hat{p}$ is well-defined. Let $Lx_1 = Lx_2$ for some $x_1, x_2 \in X$. Then $x_1 - x_2 \in N(L) \subset N(p)$ and, since p satisfies a U-condition, $p(x_1) = p(x_2)$.

This lemma together with the above remark about the properties of the inverse mapping F^{-1} form the basis of the next theorem:

Theorem 5: Let X and $\hat{X}$ be topological linear spaces and L an open, linear mapping from X onto $\hat{X}$. Assume that $p, r \in H_{\mu}(X)$ both satisfy a U-condition, that $N(L) \subset N(p) \subset N(r)$ and that r is continuous at zero in the relative topology on S(p). Unique functions $\hat{p}, \hat{r} \in H_{\mu}(\hat{X})$ then exist such that $p(x) = \hat{p}(Lx)$, $r(x) = \hat{r}(Lx)$ for all $x \in X$ and that $N(\hat{p}) \subset N(\hat{r})$. Moreover, if $\Gamma(\hat{p})$ is bounded, then $\hat{\gamma} = \gamma_{\hat{p}}(\hat{r}, \hat{X})$ and $\gamma = \gamma_p(r, X)$ are both finite and equal to each other.

Proof: Lemma 5 assures the existence of $\hat{p}, \hat{r} \in H_{\mu}(\hat{X})$ and it is readily seen that $N(\hat{p}) \subset N(\hat{r})$. But then Theorem 4 applies and states that $\hat{\gamma}$ and γ are either both infinite or both finite and $\hat{\gamma} = \gamma$. Therefore, the theorem follows directly from Theorem 1 if we can show that $\hat{r}$ is continuous at zero relative to $S(\hat{p})$. Set $I = \{t \text{ real} \mid |t| < \varepsilon\}$, then the continuity of r at the origin relative to S(p) implies that

$$Q = S(p) \cap r^{(-1)}(I) = \{x \in S(p) \mid \hat{r}(Lx) < \varepsilon\} = S(p) \cap L^{(-1)}(\hat{r}^{(-1)}(I))$$

is an open set in S(p), i.e., that $Q = S(p) \cap U$ where U is open in X. The openness of L assures that LU is open in $\hat{X}$, and now it easily follows that $LQ = LS(p) \cap LU = S(\hat{p}) \cap LU$ or that LQ is open in $S(\hat{p})$. But $LQ = \hat{r}^{(-1)}(I) \cap LS(p) = \hat{r}^{(-1)}(I) \cap S(\hat{p})$ and $\hat{r}$ is therefore continuous at zero relative to $S(\hat{p})$.

An important special case of this theorem is given in the following

Corollary 5.1: Let X be a topological linear space and $p, r \in H_{\mu}(X)$ such that $N(p) \subset N(r)$ and that both p and r satisfy U-conditions on X and are continuous at the origin. Then N(p) is a closed linear subspace of X. Consider the quotient space $\hat{X} = X/N(p)$ and the induced function $\hat{p}$ on $\hat{X}$, defined by $\hat{p}(\hat{x}) = p(x)$ where $x \in \hat{x} \in \hat{X}$. If $\Gamma(\hat{p})$ is bounded under the quotient topology on $\hat{X}$ then $\gamma_p(r,X)$ is finite.

Proof: $N(p)$ is a convex cone in X. From $\alpha p(x) \geq p(-x)$ for all $x \in X$ it follows that $N(p)$ is indeed a linear subspace of X, and since p is uniformly continuous on X, $N(p)$ is closed. This implies that the quotient map $Lx = x + N(p)$ from X onto $\hat{X}$ is linear, continuous and open. Moreover, $N(L) = N(p)$. Hence, Theorem 5 applies and the proof is complete.

This result can be combined with an argument of the type used in Corollary 1.1. To do this we begin with the following observation:

Let X be a real linear space and assume that $p \in H_\mu(X)$ satisfies a U-condition. Then the family of sets

$$\mathcal{U} = \left\{ U_t \subset X \;\middle|\; U_t = tS(p) \;,\; t > 0 \right\}$$

is already a local base of the origin for a vector topology τ_p on X. This represents a slight extension of our result on the vector topologies induced on X by symmetric and quasiconvex functions of $H_\mu(X)$. The missing condition here is the symmetry of p. As before, we see that each U_t absorbs every point of X, that $U_t \subset U_{t_1} \cap U_{t_2}$ for $t \leq \min(t_1,t_2)$, and—since p is quasiconvex of some degree α — that $U_t + U_t \subset U_{t_o}$ for $t \leq t_o/2\alpha$. Now $p(-x) \leq \alpha p(x)$ for all $x \in X$ is equivalent with $-S(p) \subset \alpha^{1/\mu} S(p)$. But then $\lambda U_t \subset U_{t_o}$ for all $-1 \leq \lambda \leq +1$ if we choose $t \leq \alpha^{-(1/\mu)} t_o$. In fact, for $0 \leq \lambda \leq 1$, obviously $\lambda\, tS(p) \subset tS(p) \subset t_o S(p)$ while for $-1 \leq \lambda < 0$, $\lambda tS(p) =$ $= |\lambda|\, t\, (-S(p)) \subset |\lambda|\, t\, \alpha^{1/\mu} S(p) \subset |\lambda|\, t_o S(p) \subset t_o S(p)$. This completes the proof that $\mathcal{U}$ is indeed a local base at zero of a vector topology on X. As before, it follows that this topology is semimetrizable.

Now we can phrase the above indicated result as follows:

Theorem 6: Let X be a complete, semimetrizable real linear space and assume that the functions $r, p \in H_{\mu}(X)$ ($p \not\equiv 0$) with $N(p) \subset N(r)$ are both lower semicontinuous on X and satisfy a U-condition on X. Then the induced map $\hat{p}$ of p on $\hat{X} = X/N(p)$ also satisfies a U-condition, and if $\hat{X}$ is complete under the $\tau_{\hat{p}}$ -topology induced by $\hat{p}$ on $\hat{X}$, then $\gamma_p(r,X)$ is finite.

Proof: Theorem 2, Part (b) assures that the functions p and r are continuous at zero, hence they are uniformly continuous on X. Then N(p) is a closed linear subspace of X and, accordingly, $\hat{X}$ is a Hausdorff space under the quotient topology $\hat{\tau}_o$ induced on $\hat{X}$ by the original topology τ_o of X. Moreover, a well-known theorem 4] states that $(\hat{X}, \hat{\tau}_o)$ is complete and metrizable. The quotient map from (X, τ_o) onto $(\hat{X}, \hat{\tau}_o)$ is continuous, open, and linear. Therefore, as in the proof of Theorem 5, it follows that on $(\hat{X}, \hat{\tau}_o)$ the induced functions $\hat{p}$ and $\hat{r}$ of p and r, respectively, are again continuous at the origin. Furthermore, it is easily checked that both $\hat{p}$ and $\hat{r}$ also satisfy U-conditions on $\hat{X}$, i.e., that $\hat{r}, \hat{p}$ are uniformly continuous on $(\hat{X}, \hat{\tau}_o)$. Consider now the vector-topology $\tau_{\hat{p}}$ induced by $\hat{p}$ on $\hat{X}$. Since $\hat{p}(\hat{x}) = 0$ implies $\hat{x} = 0$, it follows that $(\hat{X}, \tau_{\hat{p}})$ is also a Hausdorff space and hence complete and metrizable. From the continuity of $\hat{p}$ on $(\hat{X}, \hat{\tau}_o)$ it follows that the identity mapping from $(\hat{X}, \hat{\tau}_o)$ onto $(\hat{X}, \tau_{\hat{p}})$ is continuous and, of course, one-to-one, and therefore has a closed graph. But then the closed graph theorem states that the inverse is also continuous, i.e.,

---

4] The image under an open, continuous linear mapping L of a complete, semimetrizable, linear topological space X is complete and semimetrizable (see e.g. [5], p. 99).

that the topologies $\hat{\tau}_o$ and $\tau_{\hat{p}}$ are equivalent. Hence $\Gamma(\hat{p})$ is a bounded set of $(\hat{X}, \hat{\tau}_o)$ and Corollary 5.1 applies and assures the finiteness of $\gamma_p(r,X)$.

There are two possibilities of extending this theorem to complex linear spaces. The least restrictive one rests on the simple observation that in the complex case the space $(\hat{X}, \hat{\tau}_o)$ can also be considered as a real linear space and as such is still complete and metrizable; moreover, $\hat{r}$ and $\hat{p}$ remain continuous on this space. If it is now assumed in the theorem that the (real) vector topology $\tau_{\hat{p}}$ induced by $\hat{p}$ on $\hat{X}$, considered as a real linear space, is complete, one sees, as before, that the two topologies on this real space $\hat{X}$ are again equivalent. This is sufficient to assure the finiteness of $\gamma_p(r,X)$.

A different approach of extending the theorem uses the assumption that p is complex symmetric and satisfies a U-condition, i.e., that p is a seminorm on the (complex) space X. Then the proof carries over word for word. A special case of this approach is contained in the following useful Corollary:

Corollary 6.1: Let X be a complete, semimetrizable linear space and $r, p \in H_1(X)$ $(p \not\equiv 0)$ continuous seminorms on X with $N(p) \subset N(r)$. Then the induced function $\hat{p}$ of p on $\hat{X} = X/N(p)$ is a norm and if $\hat{X}$ is complete under this norm, $\gamma_p(r,X)$ is finite.

V. THE FINITE DIMENSIONAL CASE

In this section some of the earlier theorems shall be applied to the special case when X is an n-dimensional linear space. Due to the special structure of such spaces, a number of simple results about comparison-factors can be obtained. The presentation here is mainly intended to be an illustration of the earlier results; a more detailed treatment of comparison-factors on finite dimensional spaces is planned for a separate paper.

We begin with the following two basic lemmas:

Lemma 6: Let C be a convex polyhedral cone in the n-dimensional linear space X. Then every quasiconvex function $p \in H_{\mu}(C)$ is continuous (on C) at the origin, under any Hausdorff topology on X.

Proof: By assumption, there exist finitely many (non-zero) vectors $g_1, g_2, \ldots, g_m \in C$ (the extremal vectors of C) such that

$$C = \left\{ x \in X \;\middle|\; x = \sum_{k=1}^{m} \eta_k g_k, \quad \eta_k \geq 0 \right\}.$$

For $x \in C$ set $r(x) = \sum_{k=1}^{m} \eta_k$ and let q be any norm on X. Evidently $r \in H_1(C)$ and $N(r) = \{0\}$. Moreover, the set $\Gamma(r)$ is bounded, since $q(x) \leq \max_k q(g_k) \cdot r(x)$ $(x \in C)$, and as a convex polyhedron, $\Gamma(r)$ is also closed. Hence the continuous function q assumes its minimum on the compact set $\Gamma(r)$ and $q(x) \geq q(x_o)$ for all $x \in \Gamma(r)$ implies that $q(x_o) = c > 0$ because otherwise $0 \in \Gamma(r)$. Therefore $r(x) \leq \frac{1}{c} q(x)$ for all $x \in C$.

Suppose now that α is the degree of quasiconvexity of p and set $\beta = (2\alpha)^{\mu}$. Repeated application of the simple inequality $\max(\beta\max(a,b), c) \leq \beta \max(a,b,c)$ (for any real, non-negative a,b,c) then yields for all $x \in X$

$$p(x) = p\left(\sum_{k=1}^{m} \eta_k g_k\right) \leq \beta^{m} \left(\max_k p(g_k)\right)\left(\max_k \eta_k\right)^{\mu} \leq C_1 r(x)^{\mu} \leq C_1 q(x)^{\mu}$$

C_1 = constant, i.e., on C, p is continuous at zero under the norm q. This completes the proof since any Hausdorff topology on X is equivalent with the q-norm topology.

In general, the lemma does not remain valid if C is not a polyhedral convex cone. In fact, in that case, the values of q for the infinitely many extremal vectors can be arbitrarily selected without violation of any of the other assumptions of the lemma, and in particular these values can be allowed to tend toward infinity.

Lemma 7: Let C be any cone in the n-dimensional linear space X and assume that C is closed under some norm q on X. If $p \in H_\mu(C)$ is lower semicontinuous on C under q and has the property that $N(p) = \{0\}$, then S(p) is bounded under any vector topology τ on X.

Proof: Since p is lower semicontinuous on the compact set $\Gamma(q) \cap C$, it assumes its minimum on this set, i.e., there exists an $x_o \in \Gamma(q) \cap C$ such that $p(x_o) = \delta = \min (p(x) \mid x \in \Gamma(q) \cap C)$. But then $\delta > 0$, since $p(x_o) = 0$ implies $x_o = 0$ and therefore $0 \in \Gamma(q) \cap C$ which is impossible. Hence, $p(x) \geq \delta q(x)^\mu$ for all $x \in C$, or S(p) is bounded under the norm q. Since all norm topologies on X are equivalent, we also have $S(p) \subset \alpha S(q_o)$, where q_o is the ℓ_1-norm corresponding to a basis $e_1, \ldots, e_n$ of X, i.e., where $q_o(x) = \sum_{k=1}^{n} |\xi_k|$ for $x = \sum_{k=1}^{n} \xi_k e_k$. Let now U be any τ neighborhood of zero and V a balanced and absorbing neighborhood of zero such that the n-fold sum V + V + ... +V is contained in U. Then $\xi_k e_k \subset \beta V$ (k=1,..,n) for some $\beta > 0$ and hence for $x \in S(q_o)$ it follows that $x \in \beta(V+V+..+V) \subset \beta U$ or altogether $S(p) \subset \alpha S(q_o) \subset \alpha\beta U$, i.e., S(p) is bounded under τ.

There are simple counter examples which show that S(p) can be unbounded if p is not lower semicontinuous, even though $N(p) = \{0\}$. However, the condition of lower semicontinuity in Lemma 7 can be replaced by the weaker assumption that $\overline{S(p)} \subset \alpha S(p)$ for some $\alpha > 0$ where $\overline{S(p)}$ is the closure of S(p) under q. To show this, introduce for $x \in C$ the function $\bar{p}(x) = \inf(t \mid \frac{1}{t} x \in \overline{S(p)})$, then $\bar{p} \in H_1(C)$. Suppose that $x \in C$ and $\bar{p}(x) = 0$. There then exists a sequence $t_n > 0$, $\lim_{n\to\infty} t_n = 0$ such that $(1/t_n)x \in \overline{S(p)} \subset \alpha S(p)$, and hence we find that

$$p(x) = t_n^{\mu}\, p\left(\frac{1}{t_n} x\right) \leq \alpha^{\mu} t_n^{\mu}$$

or $p(x) = 0$, i.e., $x = 0$, or $N(\overline{p}) = \{0\}$. Evidently $\overline{p}$ is lower semicontinuous on C. Altogether therefore, Lemma 7 applies for $\overline{p}$; and $\overline{S(p)}$, i.e., also $S(p)$, is a bounded set under any vector topology on X.

As a direct consequence of these lemmas and of our earlier results we obtain the following

<u>Theorem 7:</u> Let C be a cone in the n-dimensional linear space X and $r, p \in H_{\mu}(C)$ $(r, p \not\equiv 0)$ such that $N(p) \subset N(r)$. Then $\gamma_p(r,C)$ is finite if one of the following conditions is satisfied:

(1) C is closed under some norm q of X, p is lower semicontinuous on C under q, $N(p) = \{0\}$, and r is continuous at zero on C under any vector topology of X.

(2) C is a polyhedral convex cone, r is quasiconvex, p is lower semicontinuous on C under some norm q, and $N(p) = \{0\}$.

(3) C is a linear subspace and both r and p satisfy (U)-conditions on C.

(1) follows from Lemma 7 and Theorem 1, Lemmas 6 and 7 together with Theorem 1 imply (2), while the same lemmas together with Theorem 6 yield (3).

We conclude this section with a simple example which contains some of the introductory results provided by M. Golomb and H. Weinberger in [3], pp 121.

Suppose that for some infinite dimensional linear space X the function $q \in H_{\mu}(X)$ satisfies a U-condition and has the property that the linear space $X_1 = N(q)$ has finite defect. Let $r, p \in H_{\mu}(X)$ $(r, p \not\equiv 0)$ be some other functions which satisfy U-conditions on X and are such that $N(p) \cap X_1 \subset N(r) \cap X_1$ and that $\gamma_o = \gamma_p(r, X_1)$ is finite. Form the new function $\hat{p} = p+q \in H_{\mu}(X)$, then $\gamma_{\hat{p}}(r,X)$ is finite.

Let X_2 be an algebraic complement of X_1 in X, i.e., assume that $X_1 \oplus X_2 = X$. By assumption, the dimension of X_2 is finite. Since $N(q) \cap X_2 = \{0\} \subset N(p) \cap X_2$ and, similarly, $N(q) \cap X_2 \subset N(r) \cap X_2$, it follows from Theorem 7 Part (3) that $p(x) \leq \gamma_1 q(x)$, $r(x) \leq \gamma_2 q(x)$ for all $x \in X_2$ and with finite γ_1 and γ_2. Let $x \in X$ be arbitrary, and $x = x_1 + x_2$, $x_1 \in X_1$, $x_2 \in X_2$, then

$$\begin{aligned} r(x) &\leq \alpha_r r(x_1) + r(x_2) \leq \alpha_r \gamma_o p(x_1) + \gamma_2 q(x_2) \leq \\ &\leq \alpha_r \alpha_p \gamma_o p(x_2) + \alpha_r \gamma_o p(x) + \gamma_2 q(x) \leq \\ &\leq (\alpha_r \alpha_p \gamma_o \gamma_1 + \gamma_2) q(x) + \alpha_r \gamma_o p(x) \leq k(q(x) + p(x)) \end{aligned}$$

where $k = \max(\alpha_r \gamma_o, \alpha_r \gamma_o \alpha_p \gamma_1 + \gamma_2)$ and α_r, α_p are the constants in the U-conditions of r and p, respectively. This completes the proof of the above statement.

VI. SARD'S QUOTIENT THEOREM AND THE HYPERCIRCLE INEQUALITY

In this Section we shall return to the special case of Example (1c). The application of the results of the previous Sections to this example provides a direct connection to important work on bounds for the error of best approximations, as developed, for example, by A. Sard [2], M. Golomb and H. Weinberger [3], P. Davis [8], and others. At the same time, this connection provides us with several possible methods for the computational evaluation of comparison-factors in this particular case.

The reason for the possibility of further development of the theory of comparison-factors in the case of Example (1c) is contained in the following simple representation theorem:

Theorem 8: Let p_X and p_Y be seminorms on the linear spaces X and Y, respectively, and let Z be a normed linear space. Consider the linear operators $A: X \to Y$, $B: X \to Z$ with domains $D_B \subset D_A \subset X$, and assume that $Bx = 0$ whenever $Ax = 0$ for $x \in D_B$. Suppose that on D_B the comparison factor $\gamma = \gamma_p(r, D_B)$ of $r(x) = \|Bx\|$ with respect to $p(x) = p_Y(Ax)$ is finite. Then there exists a linear operator $C: AX \to Z$ such that $Bx = CAx$ for all $x \in D_B$ and C is continuous with respect to p_Y and the norm on Z. Moreover,
$\gamma = \sup \left\{ \|Cy\| \;\middle|\; y \in D_B\ ,\ p_Y(y) = 1 \right\}$.

The proof is immediate. Set $Cy = Bx$ for $y = Ax \in AX \subset Y$, then C is evidently a well-defined linear operator on AX and from

$$\|Cy\| = \|Bx\| \leq \gamma\, p_Y(Ax) = \gamma\, p_Y(y) \quad , \ y \in AX$$

it follows that C is continuous on AX and that γ is the norm of C on AX.

Observe that A and B need not be continuous, but, evidently, if A is continuous then the finiteness of γ implies that B is also continuous.

The following special case of Theorem 8 was just recently discussed by M. E. Gurtin [9]: If $Z = R^1$ (the real axis), then B and C are linear functionals over D_B and AX, respectively. Hence the Hahn-Banach theorem assures the extendability of C to a linear functional $\tilde{C}$ over all of Y such that $\|\tilde{C}y\| \leq \gamma\, p_Y(y)$ for $y \in Y$.

The following example constitutes an adaption of a representation theorem of A. Sard, [2], to our discussion:

Example 5: Let $X = C^n[0,1]$ be the space of all n times continuously differentiable real functions on $I = [0,1]$ with the norm
$\|x\| = \max_k \left[\sup_{t \in I} |x^{(k)}(t)| \right]$. On the space $Y = R^n \times C^0[I]$ introduce

the norm $\|y\| = \max\left[|\eta_1|, \ldots, |\eta_n|, \sup_{t\in I}|\eta(t)|\right]$. Finally, let $Z = R^1$. For fixed $\alpha \in I$ define the linear operator $A: X \to Y$ by

$$Ax = \left\{x(\alpha), x'(\alpha), \ldots, x^{(n-1)}(\alpha), x^{(n)}(t)\right\} \in Y ;$$

then clearly $\|Ax\| \leq \|x\|$. Since $Ax = 0$ only if $x = 0$, there exists an inverse operator A^{-1}, namely

$$(14) \quad A^{-1}y \equiv x(t) = \eta_1 + \sum_{k=1}^{n-1} \frac{(t-\alpha)^k}{k!}\eta_{k+1} + \int_\alpha^t \frac{(t-s)^{n-1}}{(n-1)!}\eta(s)\, ds .$$

At the same time, this representation of A^{-1} shows that A maps X onto Y, i.e., that the domain of A^{-1} is all of Y. Moreover, the estimate $\|A^{-1}y\| \leq e\|y\|$ follows immediately from (14), which proves in turn that for the function $p(x) = \|Ax\| \in H_1(X)$ the set $S(p)$ is bounded. Let now $B: X \to Z$ be any continuous linear functional on X, and set $r(x) = \|Bx\|$. Then $N(p) = \{0\} \subset N(r)$ implies that $r(x) \leq \gamma p(x)$ and Theorem 1 assures the finiteness of $\gamma = \gamma_p(r,X)$. Hence it follows from Theorem 8 that there exists a continuous linear functional C on Y such that $Bx = CAx$ for $x \in X$. Since $R^n \oplus C^o[I] = Y$, we obtain the following well-known representation

$$Cy = \sum_{k=1}^{n} c_k\eta_k + \int_0^1 \eta(s) \quad d\lambda(s)$$

with unique coefficients $c_1,\ldots,c_n$ and a unique function λ of bounded variation on I. Altogether therefore,

$$Bx = \sum_{k=0}^{n-1} c_{k+1}x^{(k)}(\alpha) + \int_0^1 x^{(n)}(s)\; d\lambda(s)$$

In [2], A. Sard uses this result extensively to obtain error bounds for best-integration formulas. Briefly, the approach is as follows: Let $X = C^n[I]$ and $Y = C^o[I]$ with the same norms as in the example. Suppose $B: X \to R^1$ is the remainder functional of an approximate numerical integration which is exact when x is a polynomial of degree less or equal to (n-1). If $A: X \to Y$ is defined by $Ax = x^{(n)}$, then this last condition is evidently equivalent to the assumption that $Bx = 0$ whenever $Ax = 0$. Now Theorem 8 can be applied and, as in the example, the representation $Bx = \int_0^1 x^{(n)}(s)\, d\lambda(s)$ is obtained where λ is again a function of bounded variation on I. The best error bound for the numerical integration formula is the norm of B.

The assumption of the finiteness of $\gamma_p(r,X)$ used in Theorem 8 can of course be replaced by applying one of the relevant theorems of Sections III and IV. The simplest result appears to follow from an application of Corollary 6.1.

<u>Corollary 8.1:</u> Let X and Y be Banachspaces and $A: X \rightarrow Y$, $(A \not\equiv 0)$, a continuous linear operator from X <u>onto</u> Y. Suppose Z is any normed linear space and $B: X \rightarrow Z$ a continuous linear operator on X such that $Bx = 0$ whenever $Ax = 0$ $(x \in X)$. Then for the functions $p(x) = \|Ax\|$, $r(x) = \|Bx\|$, the comparison-factor $\gamma = \gamma_p(r,X)$ is finite and hence, by Theorem 8, there exists on Y a continuous linear operator $C: Y \rightarrow Z$ such that $Bx = CAx$, $(x \in X)$, and γ is the norm of C on Y.

The proof follows immediately from Corollary 6.1 if we can show that under the norm $\hat{p}$ induced by p on the quotient space $\hat{X} = X/N(p)$, this space $(\hat{X},\hat{p})$ is complete. Clearly, $\hat{X}$ is complete under the norm topology induced by the original norm on X. Moreover, by assumption, the induced linear operator $\hat{A}: \hat{X} \rightarrow Y$ maps $\hat{X}$ one-to-one and continuously onto the Banachspace Y. Hence, $\hat{A}^{-1}$ exists and is again linear and continuous. But then the norms $\hat{p}(\hat{x}) = \|\hat{A}\hat{x}\|$ and $\|\hat{x}\|$ are equivalent on $\hat{X}$, i.e., $(\hat{X},\hat{p})$ is complete.

Corollary 8.1 represents a slight generalization of Sard's "Quotient Theorem", since Z is not assumed to be complete. Of course, Corollary 8.1 could have been proved directly using the techniques of Theorem 6 which are conceptually similar to those used in [2] for the proof of the quotient theorem.

Theorem 8 provides a representation of γ as a norm of some linear operator ?. At the same time, there is a close relation between this theorem and the mapping theorems of Section 4. For the sake of simplicity, assume that $D_B = D_A = X$, then - as in Lemma 5 - we can define the mapping $F: H_1(Y) \to H_1(X)$ by setting $q(x) = \tilde{q}(Ax)$, $(x \in X)$, for every function $\tilde{q} \in H_1(Y)$. The seminorms $p, r \in H_1(X)$ of Theorem 8 both satisfy ?-conditions and - because $N(A) = N(p) \subset N(r)$ - they are therefore contained in the range of F. This means that functions $\tilde{p}, \tilde{r} \in H_1(Y)$ exist such that $p_Y(Ax) = \tilde{p}(y)$ and $\|Bx\| = \tilde{r}(y)$ where $y = Ax$. Since $\tilde{p}$ and $\tilde{r}$ are unique, we find that $\tilde{p}(y) = p_Y(y)$ and $\tilde{r}(y) = \|Cy\|$, $y \in Y$.

In the special case when $Y \subset X$ this observation leads to the well-known hypercircle inequality, which has often been used to find error bounds for best approximations, (see for example [3] and [8]).

Suppose M $(\equiv Y)$ is a linear subspace of the linear space X and M' an algebraically complementary linear subspace of M. Let $P: X \to M$ be the (unique) projection from X onto M belonging to the decomposition $M \oplus M' = X$, and $Q = I - P$ the corresponding projection from X onto M'. On X introduce some seminorm q, and as before, let Z be a normed linear space and $C: X \to Z$ a continuous linear operator. Set $r(x) = \|Cx\|$, $p(x) = q(Px)$. If $N(q) \wedge M \subset PN(C)$, it follows from Corollary 1.1 that $\gamma = \gamma_p(r, M)$ is finite and we find that $\|CPx\| \leq \gamma\, q(Px)$ for all $x \in X$. For the applications it is useful to restrict x to some linear variety $V = x_o + M$ where $x_o \in X$ is a fixed point. Then clearly $Px = x - Qx = x - Qx_o$, $(x \in V)$, and the inequality

(15) $$\|Cx - C(Qx_o)\| \leq \gamma\, q(x - Qx_o) \quad , \quad x \in V ,$$

with $\gamma = \sup \left[\|Cx\| \;\middle|\; x \in M,\ q(x) = 1 \right]$ can be considered a basic as well as general version of the hypercircle inequality.

In order to arrive at the usual form of the hypercircle inequality - as for example given in [10] , p 230 - suppose that X is a Hilbertspace, $Z = R^1$, and that M is a closed linear subspace of X. Let M' be the orthogonal complement $M^\perp$ of M, then P is the orthogonal projection onto M and P is continuous. Hence, the Pythagorean theorem applies: $\|x - Qx_o\|^2 = \|x\|^2 - \|Qx_o\|^2$, and if x is restricted to the so-called hypercircle $S_c = \{ x \in X \mid x \in V,\ \|x\| \leq c \}$, then $\|x - Qx_o\|^2 \leq (c^2 - \|Qx_o\|^2)$. Since C is now a continuous linear functional on M, we have $C(x) = (b,x)$ for all $x \in M$ where $b \in M$ is a uniquely determined point. Then $\gamma = \|b\|$ and altogether

(16) $$|Cx - C(Qx_o)| \leq \|b\| \, (c^2 - \|Qx_o\|^2)^{1/2} \quad , x \in S_c.$$

Of course, for $\|Qx_o\| > c$ the set S_c is empty and (16) vacuous. If $\|Qx_o\| \leq c$, equality holds in (16) for $z_o = Qx_o + \lambda b$ where $\lambda = (c^2 - \|Qx_o\|^2)^{1/2} / \|b\|$ when $\|b\| \neq 0$, and $\lambda = 0$ otherwise. For $b = 0$ this fact is trivial and for $b \neq 0$ it follows immediately from

$$|Cz_o - C(Qz_o)| = |\lambda| \, |Cb| = |\lambda| \, \|b\| = (c^2 - \|Qx_o\|^2)^{1/2}$$

For the case of the hypercircle inequality (16), several methods have been discussed in the literature for the computational evaluation of the comparison-factor $\gamma = \|b\|$, i.e., for the evaluation of the norm of the linear functional C on the subspace M. Basically, there are three major methods:

(a) The use of orthonormal systems , used chiefly by P. Davis [8] to compute best error bounds for a number of standard numerical integration formulas.

(b) The method of reproducing kernels , as applied to this particular problem by M. Golomb and H. Weinberger in [3].

(c) The variational approach , also discussed in [3], and applicable when X, for example, is the Hilbertspace of absolutely continuous real functions x with square-integrable derivatives over $[0,1]$ and when the inner-product has for instance the form $(x,y) = \int_0^1 (x'y' + k^2 xy)ds$.

For details, reference is made to the original publications where also other references are given.

In general, the numerical evaluation of the norm of a continuous operator C on some Banachspace presents many open questions. The most powerful methods apply only in the case when C is a linear functional and when a representation theorem is available for C. This is the particular condition used in methods (a) and (b) above, as well as in the application of Theorems 8 and 9 as developed by A. Sard in [2] . It may be interesting to investigate similar approaches when C is no longer a functional. Representation theorems are known for various types of linear operators on specific function spaces, and the application of these theorems for obtaining upper bounds for the norm of the operators may well provide a usable approach to the problem of evaluating comparison-factors in these cases.

The direct evaluation of γ - as in method (c) or related methods - appears to offer particularly challenging possibilities, but at the same time it also poses a variety of unsolved problems, especially from a numerical viewpoint. If C is a self-adjoint, completely continuous operator on a Hilbertspace, the power of the eigenvalue theory for such operators should provide particularly good methods. At the same time, it would be highly interesting to explore iterative methods which yield monotonically decreasing upper bounds for the comparison-factor.

VII. REFERENCES

[1] G.Aumann, Ueber den Vergleichsfaktor bei linearen Approximationsproblemen Numerische Mathematik 5, p68-72, 1963

[2] A.Sard, Linear Approximations Math. Surveys No. 9, Am. Mathem. Soc., Providence, R. I., 1963

[3] M.Golomb, H.Weinberger, Optimal Approximation and Error Bounds, in "On Numerical Approximation", edited by R.Langer, The Univ. of Wisconsin Press, Madison, Wisc., p117-190

[4] W.Fenchel, Convex Cones, Sets, and Functions Report of the Mathem. Dept., Princeton University Princeton, N. J., Sept. 1953

[5] J.L.Kelley, I.Namioka, and co-authors, Linear Topological Spaces, D.Van Nostrand Co., Inc., Princeton, N.J., 1963

[6] J.B.Diaz, F.T.Metcalf, Stronger Forms of a Class of Inequalities of G.Polya - G.Szego and L.V.Kantorovich Bull. Am. Math. Soc., 69, 3, p415-418, May 1963

[7] P.Henrici, Two Remarks on the Kantorovich Inequality The Am. Math. Monthly, 68, 9, p904-906, Nov. 1961

[8] P.J.Davis, Errors of Numerical Approximation for Analytic Functions, in "Survey of Numerical Analysis", edited by J.Todd, McGraw-Hill Book Company, Inc., p468-484, 1962

[9] M.E.Gurtin, A Note on Functionals Continuous with Respect
to an Operator, J. SIAM, 12, 4, p805-806, Dec. 1964
[10] P.J.Davis, Interpolation and Approximation
Blaisdell Publ. Co., 1963